

STAT100 Problem Set 5: Discrete Random Variables

You need to submit a Word document or PDF for this assignment. Make sure you do the following:

1. Upload only one file/document in ELMS for Problem Set 5
2. Include your name in the document in the upper left-hand corner. Under your name, write STAT 100 and your section number. Write Problem Set 5 centered on the page.
3. Number and letter your answers to the questions accordingly.
4. Carefully read all problems and follow all instructions.
5. Upload the assignment in ELMS before the deadline of Sunday 3/27 at 11:59 PM otherwise it is considered late. Make sure you save your document on your computer or email it to yourself so that you keep an electronic copy.

STAT100 Problem Sets need to be completed by students in RStudio. Students should refer to the Tutorial for Problem Set 5 as they are working on this problem set. All Tutorials for Problem Sets can be found in the STAT100 ELMS course under Modules.

There is no data file needed for Problem Set 5. All the necessary work can be completed in RStudio without opening a data file.

Problem Set 5 has two questions worth 25 points. Read each question carefully and follow all instructions. Please follow these instructions for providing your responses:

- For #1.a., 1.c., and 2.a., you MUST provide R code and RStudio output. In addition, you MUST include your name in the comments for the R code for #1.a., 1.c., and 2.a.
- For #1.b., 2.b., and 2.c., you can use RStudio but it is not required that you use RStudio. Be sure to *show all work in order to receive full credit*.
- For all other questions on Problem Set 5, you should type your responses directly in the document you submit for Problem Set 5.

1. (10 points) Consider a scratch ticket (scratcher) with the following probabilities of winning:

Probability of winning \$0:	$\frac{1}{3}$	Probability of winning \$5:	$\frac{5}{18}$
Probability of winning \$10:	$\frac{2}{9}$	Probability of winning \$20:	$\frac{1}{9}$
Probability of winning \$25:	$\frac{1}{18}$		

Let X be a discrete random variable defined as the prize amount for this scratch ticket. Note that all the possible prize amounts are listed above with corresponding probabilities for each possible prize amount. *HINT: See the algebraic function in the Extra Credit worksheet from Fri 3/11. The worksheet is included in the ELMS assignment for Problem Set 5.*

Answer the following questions:

- In RStudio, create a table summarizing the probability distribution for X . **You MUST use a single algebraic function for all probabilities.** See the Tutorial for Problem Set 5 for detailed instructions on how to create a table summarizing a probability distribution defined by an algebraic function. Be sure to use the appropriate names for each column of the table so that the table will be formatted in the familiar way for a probability distribution table. **In the document you upload for this assignment, for 1.a. you MUST include:**
 - the R code you used to generate the table, including comments. **IMPORTANT INSTRUCTIONS: you MUST include your name in the comments for the R code for #1.a.**
 - an image of the table from the RStudio Console
- Verify that this is a legitimate probability distribution by checking that it meets the 2 properties of probability distributions (see the Tutorial for Problem Set 5 and previously posted lecture slides). **You can use RStudio for the 2nd property, but it is not required that you use RStudio. Be sure to show all work to receive full credit.**
- In RStudio, calculate the mean of X , using the appropriate formula for the mean of a discrete random variable. See the Tutorial for Problem Set 5 for detailed instructions on how to calculate the mean of a discrete random variable in RStudio. **In the document you upload for this assignment, for 1.c. you MUST include:**
 - the R code you used to calculate the mean, including comments. **IMPORTANT INSTRUCTIONS: you MUST include your name in the comments for the R code for #1.c.**
 - an image of output from the RStudio Console showing the value of the mean
- What is $P(X > 5)$? **Do not use RStudio for 1.d.; you should type your responses directly in the document you submit for Problem Set 5. Show all work and provide your answer as a decimal rounded to three decimal places.**

2. (15 points) 33% of U.S. adults have blood type A-positive. Suppose we take a random sample of 8 U.S. adults. Let X represent the number of adults (out of the sample of 8) with blood type A-positive. This situation can be modeled with a binomial random variable X .
- In RStudio, create a table summarizing the probability distribution for X . See the Tutorial for Problem Set 5 for detailed instructions on how to create a table summarizing a probability distribution for a binomial random variable. Be sure to use the appropriate names for each column of the table so that the table will be formatted in the familiar way for a probability distribution table. **In the document you upload for this assignment, for 2.a. you MUST include:**
 - the R code you used to generate the table, including comments. **IMPORTANT INSTRUCTIONS: you MUST include your name in the comments for the R code for #2.a.**
 - an image of the table from the RStudio Console
 - Calculate the mean of this binomial random variable X , using the appropriate formula for mean of a binomial random variable. **For 2.b., you can use RStudio but it is not required that you use RStudio. Be sure to show all work to receive full credit.**
 - Calculate the standard deviation of this binomial random variable X , using the appropriate formula for standard deviation of a binomial random variable. **For 2.c., you can use RStudio but it is not required that you use RStudio. Be sure to show all work to receive full credit.**
 - What is the probability that exactly three out of the eight randomly selected adults have blood type A-positive? **Do not use RStudio for 2.d.; you should type your responses directly in the document you submit for Problem Set 5. Show all work and provide your answer as a decimal rounded to three decimal places.**
 - What is the probability that less than three out of the eight randomly selected adults have blood type A-positive? **Do not use RStudio for 2.e.; you should type your responses directly in the document you submit for Problem Set 5. Show all work and provide your answer as a decimal rounded to three decimal places.**
 - What is the probability that between two and four (2, 3, or 4) of the eight randomly selected adults have blood type A-positive? **Do not use RStudio for 2.f.; you should type your responses directly in the document you submit for Problem Set 5. Show all work and provide your answer as a decimal rounded to three decimal places.**